Analysis of the Error Performance of Trellis-Coded Modulations in Rayleigh Fading Channels

Jim Cavers and Paul Ho

CSS/LCCR TR 89-09

92-05569/

Analysis of the Error Performance of Trellis-Coded Modulations in Rayleigh Fading Channels

Jim Cavers and Paul Ho

School of Engineering Science
Simon Fraser University
Burnaby, B.C., Canada V5A 1S6

Abstract

Trellis-Coded Modulation (TCM), when combined with interleaving, is known to give good error performance in fading channels. Previously, though, the only analytical guide has been an upperbound [2], which could be very loose over the range of signal to noise ratio of interest. In contrast, this paper presents an exact expression for the pairwise error event probability of TCM transmitted over Rayleigh fading channels. It includes PSK and multilevel QAM codes, as well as coherent and partially coherent (e.g. differential, pilot tone, etc.) detection. For a coded PSK system, the exact error event probability is the upperbound of Divsalar and Simon [2] multiplied by a correction factor whose value depends on the poles of the two-sided Laplace Transform of the probability density function of the decision variable. We have found that a good estimate of the bit error probability can be obtained by considering only a small number of short error events. This implies digital computer simulation can be avoided during the link design process. We study several coded modulation schemes this way. Among the results are the fact that TCM provides significant improvment in the error floor when detected differentially, and an asymmetry in the pairwise error event probability for 16QAM.

1 INTRODUCTION

Trellis-Coded Modulations (TCM) is a class of signalling techniques which achieves good power and bandwidth efficiency [1]. Although it was originally developed for telephone channels, there has been considerable interest in recent years in applying TCM to mobile fading channels; see for example [2,3,4,5]. The principal advantage is that, when combined with interleaving of sufficient depth, TCM provides a form of time diversity, allowing the error rate to decrease with signal to noise ratio (SNR) faster than the inverse dependence commonly found on Rayleigh fading channels.

To date, the performance of TCM on fading channels has been determined by simulation, with its attendant problems of long run times and limited accuracy. An analytical expression for the error rate would be of value both as a computational tool for link design and for the insight it could provide. However, the only analytical guide has been as upperbound on the pairwise error event probability [2] :

$$\mathbf{P(i \to j)} \leq \left\{ \prod_{k \in \eta} \left(\frac{E_s}{4N_o} \right) |c_{ik} - c_{jk}|^2 \right\}^{-1}, \qquad (1)$$

where $\mathbf{C_i} = (c_{i1}, \ldots, c_{ik}, \ldots)$ is the transmitted codeword, $\mathbf{C_j} = (c_{j1}, \ldots, c_{jk}, \ldots)$ is another codeword. The c_{ik}'s and the c_{jk}'s are complex PSK symbols with a normalized magnitude equal to 1, E_s is the average received signal energy for each PSK symbol sent, and $N_o/2$ is the power spectral density (PSD) of the channel additive white Gaussian noise (AWGN). The product in the above equation is taken over the set η of k for which $c_{ik} \neq c_{jk}$. This implies if the "length" of an error event is ℓ, the upperbound will vary inversely

with $(E_s/N_o)^\ell$. What we mean by the "length" of an error event is the number of places for which the two codewords $\mathbf{C_i} = (c_{i1}, \ldots, c_{ik}, \ldots)$ and $\mathbf{C_j} = (c_{j1}, \ldots, c_{jk}, \ldots)$ are different, i.e. the Hamming distance between $\mathbf{C_i}$ and $\mathbf{C_j}$.

While the bound in (1) reveals the faster-than-inverse dependence of error rates, it is too loose over normal SNR ranges to be useful in link design. Moreover, it is restricted to trellis coded PSK, and the only cases considered were the extremes of perfect channel state information (CSI) and no CSI. In this paper, we provide an exact and easily computed expression for the pairwise error-event probability of TCM operating on a Rayleigh fading channel. It has the form of the upperbound (1) multiplied by a correction factor. The expression is quite general, including trellis coded QAM as well as PSK, and partially coherent detection, ranging from perfect CSI to pilot tones, differential detection, etc. Since the overall error rate is dominated by that of a small set of short error events, only a few terms are necessary for determination of the bit error rate to an accuracy sufficient for link analysis.

We use the result to study the behavoir of several TCM systems, both PSK and 16QAM, under coherent and partially coherent conditions.

The paper is organized as follows. We describe in Sections 2 and 3 the system model used in this study and the corresponding optimum decoder. The performance of the optimum decoder is then analyzed in Section 4. Using the results in Section 4, we provide in Section 5 the error event probabilities of various uncoded and coded modulation schemes. Also shown in Section 5 are some of our simulation results. Finally the conclusions of this study are

given in Section 6.

2 SYSTEM MODEL

The block diagram of our system is given in Fig. 1. The input to the Trellis Encoder is a sequence of binary digits denoted by $\mathbf{a} = (a_1, \ldots, a_n, \ldots)$, and the output is a sequence of coded QAM or PSK symbols denoted by $\mathbf{C_i} = (c_{i1}, \ldots, c_{ik}, \ldots)$, where $\mathbf{C_i}$ is the i^{th} codeword of the TCM scheme. Using complex notation, each c_{ik} is a point in the complex plane and the set of distinct c_{ik}'s forms the signal constellation. We normalize each signal constellation such that $\overline{|c_{ik}|^2} = 1$, where throughout this paper, the overbar is used to represent an average value or an expected value. The sequence of modulation symbols is passed to an ideal interleaver, that is, an interleaver with an infinite interleaving depth. This implies that, after deinterleaving on the receiver side, the fading process will appear independent from symbol to symbol. As shown in [6], in many cases, an interleaver with a reasonably large interleaving depth is almost as good as an ideal interleaver. The sequence of modulation symbols that appears at the output of the interleaver is denoted by $\mathbf{C'_i} = (c'_{i1}, \ldots, c'_{ik}, \ldots)$.

The baseband equivalent of the transmitted signal is

$$s(t) = A \sum_k c'_{ik} \; p(t - kT) \tag{2}$$

where $p(t)$ is the impulse response of a pulse shaping filter that satisfies Nyquist's criterion for zero intersymbol interference, $1/T$ is the symbol rate,

and A is a constant. We normalize the energy of the pulse $p(t)$ such that

$$\int_{-\infty}^{\infty} |p(t)|^2 \, dt = 1. \tag{3}$$

When $s(t)$ is transmitted over a Rayleigh fading channel, the baseband equivalent of the received signal can be written as

$$r\,(t) = g(t)s(t) + n_w(t) \tag{4}$$

where $g(t)$ is a zero mean, complex, Gaussian fading process with a variance of $\sigma_g^2 = \overline{\frac{1}{2}\,|g(t)|^2}$, and $n_w(t)$ is the baseband equivalent of the channel white Gaussian noise. The PSD of $n_w(t)$ is N_o. We assume the fading process $g(t)$ is slow enough that $g(t)$ remains roughly constant over each symbol interval. By substituting (2) into (4), and by letting g_k' denote the value of $g(t)$ during the k^{th} interval, we can rewrite the received signal in (4) as :

$$r(t) = A\left(\sum_k g_k' c_{ik}' \;\; p(t - kT)\right) + n_w(t) \tag{5}$$

The average power of the signal component of $r(t)$ is thus

$$\begin{aligned} P_s(t) &= \frac{1}{2}\overline{|g(t)s(t)|^2} \\ &= \frac{A^2}{2}\sum_\ell \sum_k \overline{c_{i\ell}'^{*} c_{ik}'}\;\overline{g_\ell'^{*} g_k'}\; p^*(t - \ell T)p(t - kT) \\ &= \frac{A^2}{2}\sum_k \overline{|c_{ik}'|^2}\;\overline{|g_k'|^2}\;|p(t - kT)|^2 \\ &= A^2\sigma_g^2 \sum_k |p(t - kT)|^2, \end{aligned} \tag{6}$$

where the asterisk denotes complex conjugate. The average received signal energy for each modulation symbol sent is thus

$$E_s \;=\; \int_0^T P_s(t)dt$$

$$= A^2\sigma_g^2 \int_{-\infty}^{\infty} |p(t)|^2\, dt$$
$$= A^2\sigma_g^2 \tag{7}$$

If each modulation symbol carries n bits of information, then the average received enery per bit is $E_b = E_s/n$.

As shown in Fig. 1, the received signal is passed to a matched filter with an impulse response equal to $p^*(t)/\sqrt{N_o}$. The output of the matched filter is sampled every T seconds to produce the sequence $\mathbf{r}' = (r'_1, \ldots, r'_k, \ldots)$ and the r'_k's are deinterleaved to produce the sequence $\mathbf{r} = (r_1, \ldots, r_k, \ldots)$. Assuming an ideal interleaver and fading slow enough to produce no appreciable intersymbol interference, r_k can be written as

$$r_k = u_k c_{ik} + n_k \tag{8}$$

with

$$u_k = \frac{A g_k}{\sqrt{N_o}}. \tag{9}$$

Just as $(r_1, \ldots, r_k, \ldots)$ denotes the deinterleaved version of $(r'_1, \ldots, r'_k, \ldots)$, the sequence $(g_1, \ldots, g_k, \ldots)$ is the deinterleaved version of the channel complex gains $(g'_1, \ldots, g'_k, \ldots)$. As such the g_k's are statistically independent and identically distributed (iid) complex Gaussian random variables, each having a zero mean and a variance of σ_g^2. This implies the u_k's in the above equations are also iid zero-mean, complex, Gaussian random variables with variance :

$$\sigma_u^2 = \frac{A^2}{N_o}\sigma_g^2 = E_s/N_o, \tag{10}$$

which also represents the signal-energy to noise-density ratio (SNR) of the system. The noise samples n_k in (8) are also iid complex Gaussian with

variance

$$\sigma_n^2 = 1. \tag{11}$$

Given that c_{ik} is transmitted, r_k is a complex Gaussian random variable with a zero mean and a variance

$$\sigma_{r|c_{ik}}^2 = \sigma_n^2 + \sigma_u^2 |c_{ik}|^2 = 1 + |c_{ik}|^2 E_s/N_o \tag{12}$$

The Channel Estimator in Fig. 1 will extract from the received signal information about the channel's complex gains. The output of the estimator/deinterleaver is the sequence $\mathbf{v} = (v_1, \ldots, v_k, \ldots)$, where v_k is the estimate of u_k and the v_k's are iid variables. The v_k's can be obtained by a variety of methods, such as a pilot tone or embedded pilot symbols, or by differential detection, in which the previous symbol in the interleaved sequence provides channel sounding energy. We assume in this study that the v_k's are complex Gaussian random variables and the covariance of u_k and v_k is $\sigma_{uv}^2 = \overline{\frac{1}{2} u_k v_k^*}$. The corresponding correlation coefficient is

$$\mu = \frac{\sigma_{uv}^2}{\sigma_u \sigma_v} \tag{13}$$

where σ_u^2 and σ_v^2 are the variances of u_k and v_k respectively. In the case of a perfect channel estimator, μ is equal to unity. As observed from (8), given that c_{ik} is transmitted, r_k and v_k are correlated Gaussian random variables with a covariance :

$$\sigma_{rv|c_{ik}}^2 = \frac{1}{2}\overline{r_k v_k^*} = c_{ik}\sigma_{uv}^2 \tag{14}$$

The two sequences $\mathbf{r} = (r_1, \ldots, r_k, \ldots)$ and $\mathbf{v} = (v_1, \ldots, v_k, \ldots)$ are the input to a Viterbi Decoder which performs Maximum Likelihood Sequence Estimation (MLSE). The output of the decoder is the sequence $\mathbf{C_j} =$

$(c_{j1}, \ldots, c_{jk}, \ldots)$, which is an estimate of the sequence $\mathbf{C_i} = (c_{i1}, \ldots, c_{ik}, \ldots)$. We will discuss more about the decoder in the next section.

3 THE OPTIMUM DECODER

Let the set of all possible codewords of the TCM scheme be denoted by

$$\{\mathbf{C}_j = (c_{j1}, \ldots, c_{jk}, \ldots)\}$$

where c_{jk} is the k^{th} symbol in the j^{th} codeword $\mathbf{C}_j$. An optimum decoder will select the codeword for which the a posteriori probability $P(\mathbf{C}_j|\mathbf{r}, \mathbf{v})$ is the largest. If all the codewords are sent with equal probability, this is equivalent to choosing the codeword with the largest conditional probability density function (pdf)

$$p(\mathbf{r}|\mathbf{C}_j, \mathbf{v}) = \prod_k p(r_k|c_{jk}, v_k), \tag{15}$$

where $p(r_k|c_{jk}, v_k)$ is the conditional probability density function of r_k given c_{jk} and v_k. Recall that u_k and v_k are correlated complex Gaussian random variables with a correlation coefficient μ, so that given v_k, u_k is a complex Gaussian random variable with a mean of $\mu v_k \sigma_u/\sigma_v$ and a variance of $(1 - |\mu|^2)\sigma_u^2$. This implies that, given v_k and c_{jk}, the random variable r_k in (8) is Gaussian with a mean of $\mu v_k c_{jk} \sigma_u/\sigma_v$ and a variance of $(1 - |\mu|^2)\sigma_u^2|c_{jk}|^2 + 1$ (recall the noise term n_k in (8) has a variance equal to unity), i.e.,

$$p(r_k|c_{jk}, v_k) = \frac{1}{2\pi\left\{\left(1 - |\mu|^2\right)\sigma_u^2|c_{jk}|^2 + 1\right\}} \exp\left[\frac{-|r_k - \mu v_k c_{jk}\sigma_u/\sigma_v|^2}{2\left\{\left(1 - |\mu|^2\right)\sigma_u^2|c_{jk}|^2 + 1\right\}}\right] \tag{16}$$

By combining the last two equations and then taking the natural log of the result, the maximum likelihood decoder is equivalent to one which selects the codeword for which the metric

$$M_j = \sum_k \left[\frac{|r_k - \mu v_k c_{jk} \sigma_u / \sigma_v|^2}{2\left\{\left(1 - |\mu|^2\right)\sigma_u^2 |c_{jk}|^2 + 1\right\}} + \ln\left\{\left(1 - |\mu|^2\right)\sigma_u^2 |c_{jk}|^2 + 1\right\} \right] \quad (17)$$

is the minimum. We assume in this study that the channel estimator is able to determine accurately the correlation coefficent μ. For convenience, we will use the following notations :

$$\beta = \mu \frac{\sigma_u}{\sigma_v} \quad (18)$$

and

$$\begin{aligned} \sigma_e^2 &= \sigma_u^2 \left(1 - |\mu|^2\right) \\ &= \left(1 - |\mu|^2\right) E_s / N_o \end{aligned} \quad (19)$$

With these notations, (17) can be written as

$$M_j = \sum_k \left[\frac{|r_k - \beta c_{jk} v_k|^2}{2\left(|c_{jk}|^2 \sigma_e^2 + 1\right)} + \ln\left\{|c_{jk}|^2 \sigma_e^2 + 1\right\} \right] \quad (20)$$

Note that if coded PSK is used, then all the $|c_{jk}|^2$'s are the same. In this case, the optimum decoding metric in (20) can be reduced to

$$M_j = \sum_k |r_k - \beta c_{jk} v_k|^2 \qquad \text{PSK only} \quad (21)$$

The metric also assumes this form if the channel estimation error is negligible compared with channel noise, that is, $|c_{jk}|^2 \sigma_e^2 \ll 1$. We will study in the next section the pairwise error event probability of the optimum decoder.

4 THE PAIRWISE ERROR EVENT PROBABILITY

Let the transmitted codeword be $\mathbf{C_i} = (c_{i1}, \ldots, c_{ik}, \ldots)$. Then according to (20), the decoder will pick the erroneous codeword $\mathbf{C_j} = (c_{j1}, \ldots, c_{jk}, \ldots)$ if

$$\sum_k \left[\frac{|r_k - \beta c_{ik} v_k|^2}{2\left(|c_{ik}|^2 \sigma_e^2 + 1\right)} + \ln\left(|c_{ik}|^2 \sigma_e^2 + 1\right) \right] > \sum_k \left[\frac{|r_k - \beta c_{jk} v_k|^2}{2\left(|c_{jk}|^2 \sigma_e^2 + 1\right)} + \ln\left(|c_{jk}|^2 \sigma_e^2 + 1\right) \right]. \quad (22)$$

By letting

$$\sigma^2[c_{ik}] = |c_{ik}|^2 \sigma_e^2 + 1, \quad (23)$$

$$\sigma^2[c_{jk}] = |c_{jk}|^2 \sigma_e^2 + 1, \quad (24)$$

and

$$\delta = \ln\left(\prod_k \frac{|c_{ik}|^2 \sigma_e^2 + 1}{|c_{jk}|^2 \sigma_e^2 + 1} \right) = \ln\left(\prod_k \frac{\sigma^2[c_{ik}]}{\sigma^2[c_{jk}]} \right), \quad (25)$$

Equ. 22 can be written as

$$\sum_k \left[\frac{|r_k|^2}{2} \left(\frac{1}{\sigma^2[c_{jk}]} - \frac{1}{\sigma^2[c_{ik}]} \right) + \frac{|v_k|^2 |\beta|^2}{2} \left(\frac{|c_{jk}|^2}{\sigma^2[c_{jk}]} - \frac{|c_{ik}|^2}{\sigma^2[c_{ik}]} \right) + 2Re\left\{ \frac{r_k^* v_k \beta}{2} \left(\frac{c_{ik}}{\sigma^2[c_{ik}]} - \frac{c_{jk}}{\sigma^2[c_{jk}]} \right) \right\} \right] \leq \delta \quad (26)$$

The left hand side of the above equation can be written as :

$$D = \sum_k D_k \quad (27)$$

where the D_k's are independent when conditioned on the codeword, and are given by

$$D_k = A_k |r_k|^2 + B_k |v_k|^2 + C_k r_k v_k^* + C_k^* r_k^* v_k, \quad (28)$$

$$A_k = \frac{1}{2}\left(\frac{1}{\sigma^2[c_{jk}]} - \frac{1}{\sigma^2[c_{ik}]}\right), \tag{29}$$

$$B_k = \frac{|\beta|^2}{2}\left(\frac{|c_{jk}|^2}{\sigma^2[c_{jk}]} - \frac{|c_{ik}|^2}{\sigma^2[c_{ik}]}\right), \tag{30}$$

and

$$C_k = \frac{\beta^*}{2}\left(\frac{c_{ik}^*}{\sigma^2[c_{ik}]} - \frac{c_{jk}^*}{\sigma^2[c_{jk}]}\right) \tag{31}$$

Note that the expressions simplify significantly in the case of PSK, since the A_k's, B_k's, and δ are all zero. The probability of confusing the codeword $\mathrm{C_i} = (c_{i1}, \ldots, c_{ik}, \ldots)$ with the codeword $\mathrm{C_j} = (c_{j1}, \ldots, c_{jk}, \ldots)$ is

$$\mathrm{P(i \rightarrow j)} = Prob(D < \delta) = \int_{-\infty}^{\delta} p(D)dD \tag{32}$$

where $p(D)$ is the pdf of the random variable D. We will show next how to evaluate this error probability.

As mentioned earlier, we assume in this study that v_k is a Gaussian random variable. Also, it is observed from (8) that, given c_{ik} is transmitted, r_k is a Gaussian random variable. In other words, the variable D_k in (28) is a quadratic form of zero mean Gaussian variates. Let $\phi_k(s)$ be the two-sided Laplace transform of the pdf of the random variable D_k. As shown in [7, (4B.5)], as well as in [8, (B-3-16)], $\phi_k(s)$ is

$$\phi_k(s) = \frac{p_{1k}p_{2k}}{(s - p_{1k})(s - p_{2k})} \tag{33}$$

where

$$p_{1k} = w_k - \sqrt{w_k{}^2 - \frac{1}{4x_k y_k}}, \tag{34}$$

$$p_{2k} = w_k + \sqrt{w_k{}^2 - \frac{1}{4x_k y_k}}, \tag{35}$$

are the left half plane and right half plane poles, respectively, and

$$w_k = -\frac{A_k\sigma^2_{r|c_{ik}} + B_k\sigma^2_v + C_k\sigma^2_{rv|c_{ik}} + C_k^*\sigma^2_{rv|c_{ik}}^*}{4x_k y_k}, \tag{36}$$

$$x_k = \sigma^2_{r|c_{ik}}\sigma^2_v - \left|\sigma^2_{rv|c_{ik}}\right|^2, \tag{37}$$

$$y_k = A_k B_k - |C_k|^2. \tag{38}$$

Note that the region of convergence for $\phi_k(s)$ in (33) is the vertical strip containing the imaginary axis but not including either pole.

Because of the independence of the D_k's, the two-sided Laplace Transform of $D = \sum_k D_k$ is

$$\begin{aligned}\Phi_D(s) &= \prod_{k\in\eta}\phi_k(s) \\ &= \prod_{k\in\eta}\frac{p_{1k}p_{2k}}{(s-p_{1k})(s-p_{2k})},\end{aligned} \tag{39}$$

where the product is taken over the set η of k for which $c_{ik} \neq c_{jk}$ and the region of convergence is the vertical strip enclosing the $j\omega$ axis bounded by the closest poles on either side. As pointed out earlier, if $c_{ik} = c_{jk}$, the term D_k is zero. The pdf of D is

$$p(D) = \zeta^{-1}\left\{\Phi_D(s)\right\} \tag{40}$$

where $\zeta^{-1}(\bullet)$ denotes the inverse Laplace Transform operation. At this point, we could invert (39) by residues or partial fraction expansion, and perform the integration (32) to obtain the error probability. However, it is simpler to express the error event probability as :

$$\mathbf{P(i \rightarrow j)} = Prob(D < \delta) = F_D(\delta) \tag{41}$$

where the cumulative density function $F_D(\bullet)$ has a Laplace transform

$$\Psi_D(s) = \frac{\Phi_D(s)}{s} \tag{42}$$

The pairwise error event probability can therefore be written as

$$\mathbf{P(i \to j)} = \begin{cases} -\sum \text{Residue}\left[e^{s\delta}\Phi_D(s)/s\right]_{RHP poles}, & \delta \le 0 \\ \sum \text{Residue}\left[e^{s\delta}\Phi_D(s)/s\right]_{LHP poles}, & \delta > 0 \end{cases} \tag{43}$$

which is both exact and easily computed, especially in the case of simple poles.

In the case of trellis codes designed by Set Partitioning [1], the minimum Euclidean distance error events have multiple poles, which somewhat complicates the calculation of the residues in (43). If the multiplicity is small, differentiation by hand is straightforward. A more mechanical procedure is desirable, though, and for this we can resort to numerical contour integration, or to an FFT method [10].

We have achieved the goal of an exact expression for the pairwise error event probability. However (43), along with other equations, do not provide enough insight into the important factors that govern the error performance of a coded system. For this reason, we will simplify these equations as much as possible.

From (34) and (35), it is clear that $p_{1k}p_{2k} = (4x_k y_k)^{-1}$. Then using (12), (13), (14), (18), (23), (24), (29), (30), (31), (37), and (38), one can simplify $p_{1k}p_{2k}$ to :

$$p_{1k}p_{2k} = -\frac{\left\{1 + |c_{jk}|^2(1 - |\mu|^2)E_s/N_o\right\}}{|\mu|^2 d_k^2 E_s/N_o}, \tag{44}$$

where

$$d_k^2 = |c_{ik} - c_{jk}|^2 \tag{45}$$

is the squared Euclidean distance between c_{ik} and c_{jk} in the signal space. Using the same set of equations mentioned above, one can show that

$$w_k = -\frac{(1-|\mu|^2)(|c_{jk}|^2 - |c_{ik}|^2) - |\mu|^2 d_k^2}{2|\mu|^2 d_k^2}. \tag{46}$$

Substituting (44) into (39), we have

$$\Phi_D(s) = \left(\prod_{k \in \eta} \frac{\left\{1 + |c_{jk}|^2 (1-|\mu|^2) E_s/N_o\right\}}{|\mu|^2 d_k^2 E_s/N_o} \right) \left(\prod_{k \in \eta} \frac{-1}{(s-p_{1k})(s-p_{2k})} \right) \tag{47}$$

where p_{1k} and p_{2k} can be rewritten as :

$$p_{1k} = w_k - \sqrt{w_k{}^2 - p_{1k}p_{2k}} \tag{48}$$

$$p_{2k} = w_k + \sqrt{w_k{}^2 - p_{1k}p_{2k}} \tag{49}$$

Note that in the case of coded PSK and perfect channel state information, $|c_{ik}|^2 = |c_{jk}|^2 = 1$, and $|\mu|^2 = 1$. This implies (25), (44), (46) and (47) become :

$$\delta = 0, \qquad \text{PSK} \tag{50}$$

$$p_{1k}p_{2k} = -\left\{ \left(\frac{E_s}{N_o} \right) d_k^2 \right\}^{-1}, \qquad \text{PSK, Perfect CSI} \tag{51}$$

$$w_k = \frac{1}{2} \qquad \text{PSK} \tag{52}$$

and

$$\Phi_D(S) = \left\{ \prod_{k \in \eta} \left(\frac{E_s}{4N_o} \right) |c_{ik} - c_{jk}|^2 \right\}^{-1} \left(\prod_{k \in \eta} \frac{-1}{4(s-p_{1k})(s-p_{2k})} \right)$$
$$\text{PSK, Perfect CSI} \tag{53}$$

respectively. **Note that the first product term in (53) is the upperbound in (1).** In otherwords, the pairwise error event probability is the upperbound in (1) multiplied by a correction factor whose value depends on poles of the Laplace Transform of the pdf $p(D)$.

As in (1), (47) and (53) suggest that at large enough signal to noise ratio, the average error probability of the coded modulation scheme will be dominated by the shortest error events and the error curve will vary as $(E_s/N_o)^{-L_{min}}$, where L_{min} is the length of the shortest error events. **In other words, the code diversity is equal to the "length" of the shortest error events of the underlying TCM scheme.** Thus in designing a coded system for Rayleigh fading channels, one should choose a TCM scheme with as large a value of L_{min} as the computational complexity of the decoder allows. A rule of thumb is that the decoding complexity increases exponentially with L_{min}. It should be pointed out that the shortest error events are not necessarily the minimum distance error events. For example, the code in Fig. 2 has a minimum squared Euclidean distance of 4.585 and the length of the minimum distance events is 3. On the other hand, the shortest error events have length $L_{min} = 2$ and the corresponding squared distance is 6.

In most digital communication systems, the average bit error probability is of greater interest than the pairwise error event probability. A good estimate of the bit error probability can be obtained by considering only a small set of short error events as follows :

$$\mathrm{P_b} \approx \frac{1}{n} \sum_j m_{ij} \mathrm{P(i \rightarrow j)}, \tag{54}$$

with

$$L_{min} \leq d_H(\mathbf{C_i}, \mathbf{C_j}) \leq L_{min} + N. \tag{55}$$

In the above equations, n is the number of input bit per encoding interval, m_{ij} is the number of bit-error associated with each error event, $d_H(\mathbf{C_i}, \mathbf{C_j})$ is the Hamming distance or the length of the error event between $\mathbf{C_i}$ and $\mathbf{C_j}$, L_{min} is the length of the shortest error events, and N is usually a small integer. For PSK systems, we assume the bit error probability is always independent of the codeword sent. Therefore, $\mathbf{C_i}$ is chosen to be the all zero-phase codeword when we calculate the bit error probability of the TCM scheme in Fig. 2. It should be pointed out that the assumption about the bit error probability being independent of the codeword sent is not valid for QAM systems. This point will be explained in details in Section 5.

In summary, (25), and (43) to (49) are all the equations required to calculate the pairwise error event probability of TCM. In the next section, we will illustrate the use of these equations through several examples.

5 EXAMPLES

We provide in this section the pairwise error event probabilities of various uncoded and trellis-coded modulation schemes. These are (i) uncoded BPSK with perfect channel state information, (ii) trellis-coded PSK with perfect channel state information, (iii) uncoded and trellis-coded PSK with differential detection, and (iv) trellis-coded PSK and QAM with a feedforward pilot tone for channel state estimation. Once again, we point out that these

results are valid only for the case of ideal interleaving.

5.1 Uncoded BPSK

As an example of the use of (43), we obtain the well known result for uncoded BPSK with perfect channel state information. Dropping the subscript k in (48) to (53) and using the fact that $c_i = 1$ and $c_j = -1$ (or vice versa), we have

$$\Phi_D(s) = \left(\frac{E_s}{N_o}\right)^{-1} \frac{-1}{4(s-p_1)(s-p_2)} \tag{56}$$

where

$$p_1 = \frac{1}{2}\left(1 - \sqrt{1 + \left(\frac{E_s}{N_o}\right)^{-1}}\right), \tag{57}$$

and

$$p_2 = \frac{1}{2}\left(1 + \sqrt{1 + \left(\frac{E_s}{N_o}\right)^{-1}}\right) \tag{58}$$

Substituting these equations, as well as $\delta = 0$, into (43) yields

$$\begin{aligned} \mathbf{P_b} &= -\text{Residue}\left[\Phi_D(s)/s\right]_{p_2} \\ &= \frac{1}{2}\left(\frac{E_s}{N_o}\right)^{-1} \frac{1}{1 + (\frac{E_s}{N_o})^{-1} + \sqrt{1 + (\frac{E_s}{N_o})^{-1}}} \\ &= \frac{1}{2}\left(1 - \sqrt{\frac{E_s/N_o}{1 + E_s/N_o}}\right) \end{aligned} \tag{59}$$

which is identical to (7.3.7) in [7].

5.2 Trellis-Coded PSK

Consider the trellis-coded 8PSK (octal PSK) scheme shown in Fig. 2. As shown in [1], this code has a minimum squared Euclidean distance of 4.585 and the length of the minimum distance events is 3. From the trellis diagram, we can see the length of the shortest error events is 2. Example of a shortest error event is depicted by shading in Fig. 2. This corresponds to the case when the transmitted word is

$$\mathbf{C_i} = (1, 1, \ldots)$$

and the erroneous word is

$$\mathbf{C_j} = (e^{j2\pi/4}, e^{j4\pi/4}, \ldots),$$

where $e^{jn\pi/4}$ is the complex value for the n^{th} signal point. The squared Euclidean distance between these two codewords is $d_1^2 + d_2^2 = 2 + 4 = 6$. Assuming we have perfect channel state information, then $\Phi_D(s)$ in (53) becomes

$$\Phi_D(s) = \frac{1}{8}\left(\frac{E_s}{N_o}\right)^{-2} \frac{1}{(s - p_{11})(s - p_{21})(s - p_{12})(s - p_{22})}, \tag{60}$$

where

$$\begin{bmatrix} p_{11} \\ p_{21} \end{bmatrix} = \frac{1}{2}\left(1 \mp \sqrt{1 + 2\left(\frac{E_s}{N_o}\right)^{-1}}\right) \tag{61}$$

and

$$\begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix} = \frac{1}{2}\left(1 \mp \sqrt{1 + \left(\frac{E_s}{N_o}\right)^{-1}}\right) \tag{62}$$

The corresponding pairwise error event probability is

$$\begin{aligned} \mathbf{P(i \to j)} &= -\text{Residue}\left[\Phi_D(s)/s\right]_{p_{21}} - \text{Residue}\left[\Phi_D(s)/s\right]_{p_{22}} \\ &= \frac{1}{\Gamma^2} \frac{1+\Gamma\left(\sqrt{1+2\Gamma^{-1}} - \sqrt{1+\Gamma^{-1}}\right)}{\left(1+\sqrt{1+2\Gamma^{-1}}\right)\left(1+\sqrt{1+\Gamma^{-1}}\right)\left(\sqrt{1+2\Gamma^{-1}}\right)\left(\sqrt{1+\Gamma^{-1}}\right)} \end{aligned} \tag{63}$$

where $\Gamma = E_s/N_o$. In contrast, the upperbound in (1) for the shortest error event is $\mathbf{P(i \to j)} \leq 2\Gamma^{-2}$. The exact error probability and the upperbound, as a function of the signal to noise ratio E_b/N_o, are shown in Fig. 3. Note the energy per bit, E_b, is equal to $E_s/2$, since each PSK pulse sent carries two bits of information. It is observed from Fig. 3 that at large signal to noise ratio, the upperbound is about 4 dB away from the exact value !

As mentioned earlier, in many cases the bit error probability is of greater interest than the pairwise error event probability. We obtain an estimate of the bit error probability by first identifying all error events of length up to 4 (relative to the all zero-phase codeword in Fig. 2), calculate the exact bit error probability of each of these error events, and adding these bit error probabilities to obtain the estimate. The results are shown in Fig. 4. Also shown in Fig. 4 are our simulation results. It is observed that the analytical estimates and the simulation results are very close, especially at large signal to noise ratio. This suggests time consuming simulations can be avoided during the link design process.

5.3 Uncoded and Trellis-Coded DPSK

In a DPSK (Differential PSK) system, the sequence $\mathbf{C_i} = (c_{i1}, \ldots, c_{ik}, \ldots)$ that appears at the output of the trellis encoder will be differentially encoded into another sequence denoted by $\mathbf{z} = (z_1, \ldots, z_k, \ldots)$, where $z_k = c_{ik} z_{k-1}$. This implies the term c_{ik} in (8) should now be replaced by $c_{ik} z_{k-1}$ and the term u_k replaced by $A g_k z_{k-1}/\sqrt{N_o}$. At the receiver, the differential detection process can be represented by setting

$$v_k = r_{k-1} = \frac{A g_{k-1}}{\sqrt{N_o}} z_{k-1} + n_{k-1}. \tag{64}$$

Since the noise term n_{k-1} has a variance of 1 (see (11)) and the variance of g_{k-1} is E_s/N_o (see (10)), this implies

$$\sigma_v^2 = 1 + \frac{E_s}{N_o}, \tag{65}$$

By combining (64) and (9), we have

$$\sigma_{uv}^2 = \frac{1}{2}\overline{u_k v_k^*} = \frac{A^2}{N_o} R_g(T), \tag{66}$$

where $R_g(\tau)$ is the autocorrelation function of the fading process. For a land mobile radio system, the autocorrelation function can be modelled as [9, (6-108)]

$$R_g(\tau) = \sigma_g^2 J_o(2\pi f_d \tau) \tag{67}$$

where $J_o(\bullet)$ is the Bessel function of order zero and f_d is the maximum Doppler frequency. Substituting (10), (65), (66),(67) into (13) we have

$$|\mu|^2 = \frac{J_o^2(2\pi f_d T)}{1 + \left(\frac{E_s}{N_o}\right)^{-1}} \tag{68}$$

The term $|\mu|^2$can then be used in (44) to (49) to obtain $\phi_D(S)$.

In the case of uncoded binary DPSK, (53) becomes

$$\Phi_D(s) = \frac{1+(1-|\mu|^2)(E_s/N_o)}{4|\mu|^2(E_s/N_o)} \frac{-1}{(s-p_1)(s-p_2)}, \tag{69}$$

where

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \frac{1}{2}\left(1 \mp \sqrt{\frac{1+E_s/N_o}{|\mu|^2 E_s/N_o}}\right) \tag{70}$$

The corresponding bit error probability (43) is

$$\begin{aligned} \mathbf{P_b} &= -\text{Residue}\left[\Phi_D(s)/s\right]_{p_2} \\ &= \frac{1+(1-|\mu|^2)E_s/N_o}{2|\mu|^2 E_s/N_o} \frac{1}{\left(1+\sqrt{\frac{1+E_s/N_o}{|\mu|^2 E_s/N_o}}\right)} \frac{1}{\left(\sqrt{\frac{1+E_s/N_o}{|\mu|^2 E_s/N_o}}\right)} \\ &= \frac{1}{2}\left(1-\sqrt{\frac{|\mu|^2 E_s/N_o}{1+E_s/N_o}}\right) \end{aligned} \tag{71}$$

In the case of slow fading, i.e., when $f_dT = 0$, the term $|\mu|^2$ in (68) is roughly equal to $\{1+(E_s/N_o)^{-1}\}^{-1}$ and $\mathbf{P_b}$ becomes $2/(1+E_s/N_o)$, which is identical to (7.3.10) in [7].

In the case when $E_s/N_o \to \infty$, $|\mu|^2 \approx J_o^2(2\pi f_dT)$ and the bit error probability becomes

$$\begin{aligned} \mathbf{P_b} &\approx \frac{1}{2}(1-|\mu|) \\ &\approx \frac{1}{2}(1-J_o(2\pi f_dT)) \qquad E_s/N_o \to \infty \end{aligned} \tag{72}$$

The last equation gives the irreducible error probability of binary DPSK as a function of the maximum Doppler frequency. For very fast fading,

the Bessel function approaches zero and the irreducible error probability becomes one half. On the other hand, for slow fading, i.e., when $f_dT \approx 0$, $J_o(2\pi f_dT) \approx 1-(\pi f_dT)^2$, and

$$\mathbf{P_b} \approx \frac{1}{2}(\pi f_dT)^2, \qquad E_s/N_o \to \infty, f_dT \approx 0. \tag{73}$$

Let us consider next the case of trellis-coded DPSK. We will once again use the code in Fig. 2 as an example. As before, we consider the shortest error event associated with the two codewords $\mathbf{C_i} = (1,1,\ldots)$ and $\mathbf{C_j} = (e^{j2\pi/4}, e^{j4\pi/4}, \ldots)$. Since the squared Euclidean distance between these two codewords is $d_1^2 + d_2^2 = 2+4 = 6$, the Laplace Transform corresponding to this error event is

$$\Phi_D(s) = \frac{1}{8}\left(\frac{1+(1-|\mu|^2)(E_s/N_o)}{|\mu|^2(E_s/N_o)}\right)^2 \frac{1}{(s-p_{11})(s-p_{21})(s-p_{12})(s-p_{22})}, \tag{74}$$

where

$$\begin{bmatrix} p_{11} \\ p_{21} \end{bmatrix} = \frac{1}{2}\left(1 \mp \sqrt{\frac{2+(2-|\mu|^2)(E_s/N_o)}{|\mu|^2(E_s/N_o)}}\right) \tag{75}$$

and

$$\begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix} = \frac{1}{2}\left(1 \mp \sqrt{\frac{1+(E_s/N_o)}{|\mu|^2(E_s/N_o)}}\right) \tag{76}$$

The corresponding pairwise error event probability is

$$\begin{aligned} \mathbf{P(i \to j)} &= -\text{Residue}\left[\Phi_D(s)/s\right]_{p_{21}} - \text{Residue}\left[\Phi_D(s)/s\right]_{p_{22}} \\ &= \frac{1}{\Gamma_e^2}\frac{1+\Gamma_e\left(\sqrt{1+2\Gamma_e^{-1}}-\sqrt{1+\Gamma_e^{-1}}\right)}{\left(1+\sqrt{1+2\Gamma_e^{-1}}\right)\left(1+\sqrt{1+\Gamma_e^{-1}}\right)\left(\sqrt{1+2\Gamma_e^{-1}}\right)\left(\sqrt{1+\Gamma_e^{-1}}\right)}, \end{aligned} \tag{77}$$

where the effective signal to noise ratio is

$$\Gamma_e = \frac{|\mu|^2 E_s/N_o}{1+(1-|\mu|^2)E_s/N_o} \tag{78}$$

For large E_s/N_o, $\Gamma_e \approx |\mu|^2/(1-|\mu|^2)$ and the error probability simplifies to

$$\mathbf{P(i \to j)} \approx (1-|\mu|)\left(1-\frac{1-|\mu|}{\left(|\mu|+\sqrt{2-|\mu|^2}\right)\sqrt{2-|\mu|^2}}\right) \qquad E_s/N_o \to \infty \tag{79}$$

which is the irreducible error rate of this particular error event. Note that $|\mu| \approx J_o(2\pi f_d T)$ when the signal to noise ratio approaches infinity. For very fast fading, μ approaches zero and the irreducible error probability approaches 1/2. On the otherhand, for very slow fading ($f_d T \approx 0$), μ approaches unity and the irreducible error probability becomes :

$$\begin{aligned} \mathbf{P(i \to j)} &\approx \frac{3}{2}(1-|\mu|)^2 \\ &\approx \frac{3}{2}(1-J_o(2\pi f_d T))^2 \\ &\approx \frac{3}{2}(\pi f_d T)^4 \qquad E_s/N_o \to \infty, f_d T \approx 0 \end{aligned} \tag{80}$$

Comparing (73) and (80), we conclude the diversity effect introduced by coding applies to the irreducible error floor as well, since it is lowered by a factor of $3(\pi f_d T)^2$.

We show in Fig. 5 the bit error probability of this coded DPSK scheme. As in the case of coherent detection, we consider only error events of length up to 4 in our bit error probability calculation. Also shown in Fig. 5 are the simulation results taken from [6]. It is observed that for normalized fade rates of .01 and .03, the simulation results agree quite well with the analytical

estimates, especially at large signal to noise ratio. For purpose of comparison, we include in Fig. 5 the coherent PSK results from Fig. 4. It is observed that for a normalized fade rate of .01, coded DPSK requires three to five dB more energy to achieve the same error performance as coherent coded PSK. For higher fade rates, the performance of coded DPSK deteriorates rapidly, as the error floor becomes significant. Recall, though, that at a normalized fade rate of .03, the error floor in uncoded DPSK is larger by a factor of 22.5. The advantage of DPSK, of course, is that no complicated phase recovery circuit is required.

5.4 PSK/QAM with a Pilot Tone

In order to have an accurate channel estimator at the receiver, a pilot signal can be sent along with the data bearing signal in (2). The pilot signal can be a tone (or multiple tones) or it can be a known sequence of symbols that are inserted periodically into the data bearing signal. We analyze in this section the performances of coded modulation schemes with a reference pilot tone.

With a pilot tone, the transmitted signal in (2) becomes

$$s(t) = \alpha \ + \ A\left(\sum_k c'_{ik} \ p(t-kT)\right) \tag{81}$$

where α is the amplitude of the pilot signal. The corresponding received signal (assuming the fading process is constant over each interval) is

$$r(t) = \alpha\left(\sum_k g'_k \ w(t-kT)\right) + A\left(\sum_k g'_k c'_{ik} \ p(t-kT)\right) + n_w(t), \tag{82}$$

which is identical to (4) except for an additional term representing the received pilot tone. Note that in the above equation $w(t) = 1$ for $0 \leq t \leq T$ and $w(t) = 0$ otherwise. It is assumed the power spectrum of the data bearing signal has a spectral null that allows for the insertion and the extraction of the pilot.

From the first equation we see the fraction of the total energy spent on transmitting the data signal and the pilot signal is $1/(1+r)$ and $r/(1+r)$ respectively, where

$$r = \frac{\alpha^2 T}{A^2}. \tag{83}$$

Recall the term E_s is the received signal energy per symbol of a non pilot-aided detection system. This implies if the same amount of energy is used in a pilot-aided detection system, the received data signal energy is only $E_s/(1+r)$. **Therefore, for a pilot-aided detection system, we should replace the term E_s in all previous equations by $E_s/(1+r)$.**

At the receiver side, the channel estimator in Fig. 1 is simply a pilot extraction filter with a frequency response

$$H(f) = \begin{cases} \frac{1}{\sqrt{N_o}} & -B_p/2 \leq f \leq B_p/2 \\ 0 & \text{otherwise} \end{cases} \tag{84}$$

Note that the bandwidth B_p of the filter must be wide enough to allow the fading process to pass through undistorted. This implies $B_p \geq 2f_d$, where f_d is the maximum Doppler frequency. The output of the channel estimator/de-interleaver is

$$v_k = \frac{\alpha g_k}{\sqrt{N_o}} + \eta_k \tag{85}$$

where η_k is a zero mean Gaussian noise term with a variance of B_p. Note that η_k is independent of the noise term n_k in (8). Because of the extra

energy spent on the pilot signal, the term $A^2\sigma_g^2$ in (7) becomes $E_s/(1+r)$. This implies the variance of v_k is

$$\sigma_v^2 = B_p + \frac{r}{1+r}\frac{1}{T}\frac{E_s}{N_o} \tag{86}$$

Combining (85) and (9), we have the covariance

$$\sigma_{uv}^2 = \frac{E_s}{N_o}\frac{\sqrt{r/T}}{1+r} \tag{87}$$

With the scaling factor of $(1+r)$, the variance of u_k in (10) becomes $(E_s/N_c)/(1+r)$. This implies

$$|\mu|^2 = \frac{1}{1+\left(\frac{1+r}{r}\right)(B_pT)\left(\frac{E_s}{N_o}\right)^{-1}} \tag{88}$$

The above equation can then be substituted into (40) to (49) for calculating the pairwise error event probability. As pointed out earlier, the term E_s/N_o in (40) to (49) must be replaced by $(E_s/N_o)/(1+r)$ because of the extra energy spent in sending the pilot tone.

The analytical results for the eight state code, with pilot-aided detection, are shown in Fig. 6. Also shown are the analytical results of DPSK taken from Fig. 5. Three different bandwidths of the pilot filter, $B_pT = .02, .06, .12$, corresponding to a maximum f_dT of .01, .03, .06, are considered in Fig. 6. Note that for each value of B_pT, there exists an optimum value for the energy ratio r. We found that for $B_pT = .02, .06$, and .12, the optimum energy ratio r is (approximately) .1, .25, and .3 respectively. It is observed from Fig. 6 that for any fade rate smaller than .06, the performance of this coded modulation scheme, with an appropiate choice of the bandwidth of the pilot filter, is quite close to the idealized result from Fig. 4 (the result in

Fig. 4 is based on the assumption that no pilot signal has been sent and yet the receiver is able to extract perfect channel information from the received signal).

Finally let us consider uncoded 16 QAM with a pilot tone. The signal constellation for 16 QAM is shown in Fig. 7. Each signal point is a complex number of the form $(n\Delta, m\Delta)$, where $n, m \in \{\pm 1, \pm 3\}$. Since we assume the mean square value of the signal points is equal to unity, the parameter Δ in Fig.7 is equal to $1/\sqrt{10}$. Because the signal points have different magnitude, the parameter δ in (25) is not necessarily zero. As shown in (43), when $\delta \leq 0$, the error event probability is the sum of residues at the poles on the right hand plane. On the other hand, when $\delta > 0$, the poles on the left hand plane should be used instead. Another consequence of having signal points with different magnitudes is that the Laplace transform in (47) will be different for different error events that have the same squared Euclidean distance. This suggests the pairwise error event probability exhibits an asymmetric property, i.e., the probability of confusing the transmitted symbol X with the symbol Y can be different from the probability of confusing the transmitted symbol Y with the symbol X. We show in Fig. 8 the probability of confusing the signal point

1. $c_i = (\Delta, 3\Delta)$ with $c_j = (3\Delta, 3\Delta)$,
2. $c_i = (3\Delta, 3\Delta)$ with $c_j = (\Delta, 3\Delta)$,
3. $c_i = (\Delta, \Delta)$ with $c_j = (\Delta, -\Delta)$,

when the pilot filter has a bandwidth of $B_pT = .12$, and the energy ratio is

$r = .25$. The symbols in all these error events are separated by the minimum squared Euclidean distance of 0.4. The second error event is the first event with the transmitted and decoded symbols reversed. It is the worst case error event for 16 QAM, due to the facts that c_i and c_j are separated by the minimum distance, and that the larger the transmitted signal's amplitude (the point $(3\Delta, 3\Delta)$ has the largest possible amplitude), the larger the variance of the received sample (see (12) and (16)) and hence a higher chance of making an error. In general, for the same squared Euclidean distance, the probability is higher for error events between symbols with larger magnitudes than between symbols with smaller magnitudes. When we have extremely slow fading, we can, in principle, set the bandwidth of the pilot filter to zero and the term $|\mu|^2$ in (88) becomes unity, independent of the signal to noise ratio. In this case, the asymmetric property of the pairwise error event probability disappears.

6 CONCLUSIONS

We have derived in this paper an exact and easily computed expression for the pairwise error event probability for trellis-coded modulations transmitted over Rayleigh fading channels. For coded PSK schemes with perfect channel state information at the receiver, our expression for the pairwise error event probability is the loose upperbound of Divsalar and Simon [2] multiplied by a correction factor whose value depends on the poles of the two-sided Laplace Transform of the probability density function of the decision variable. However, our results are not limited to perfect channel state information,

nor to PSK. By considering only a small set of short error events, we obtain a good estimate of the bit error probability of various coded modulation schemes. This implies that time consuming digital computer simulations can be avoided during the link design process.

We found that with differential detection, Ungerboeck's 8 state 8PSK code can perform quite well up to the very high normalized fade rate of 0.06. Beyond this value, the differential phase jitter introduced by fading causes the error floor to exceed 10^{-2}, a common definition of threshold. The irreducible error floors is nevertheless significantly lower than that for uncoded DPSK, because of the time diversity effect of interleaved TCM. We found out for high fades rates, partial coherent PSK, with a pilot tone inserted for providing channel state information, can be used to eliminate the problem of irreducible error floors associated with differential detection.

We have studied the performance of uncoded 16 QAM in a Rayleigh fading channel. We found that the pairwise error event probability exhibits an asymmetric property, i.e. the probability the decoder chooses the symbol X given Y is transmitted is different from the probability of picking Y given X is transmitted. Although we have only considered the case of Rayleigh fading, our results can be easily extended to the more general case of Rician fading.

References

[1] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signal",

IEEE Trans. on Information Theory, Vol. IT-28, no. 1, pp. 55-66, January 1982.

[2] D. Divsalar and M.K. Simon, **"The Design of Trellis Coded MPSK for Fading Channels : Performance Criteria"**, *IEEE Trans. on Communications*, Vol. 36, no. 9, pp. 1004-1012, September 1988.

[3] D. Divsalar and M.K. Simon, **"The Design of Trellis Coded MPSK for Fading Channels :Set Partitioning for Optimum Code Design"**, *IEEE Trans. on Communications*, Vol. 36, no. 9, pp. 1013-1021, September 1988.

[4] P.J. McLane, P.H. Wittke, P.Ho, and C. Loo, **"The Performance of Trellis Codes for Mobile Satellite Communications"**, *IEEE Trans. on Communications*, Vol. 36, no. 9, pp. 1242-1246, November 1988.

[5] S.G. Wilson and Y.S. Leung, **"Trellis-Coded Phase Modulation on Rayleigh Channels"**, *Proc., IEEE International Conf. on Communications*87, Seattle, WA., June 7-10, 1987.

[6] Y. Wu and P. Ho, **"The Performance of Trellis-Coded DPSK for Rayleigh Fading Channels"**, Proc., *IEEE Pacific Rim Conference on Communications*, Victoria, B.C., June 1989.

[7] J.G. Proakis, **Digital Communications**, *McGraw Hill*, 1983.

[8] M. Schwartz, W. Bennet, and S. Stein, **Communication Systems and Techniques**, *McGraw Hill*, 1966.

[9] W.C.Y. Lee, **Mobile Communications Engieering**, *McGraw Hill*, 1982.

[10] J.K. Cavers, **"On the Fast Fourier Transform Inversion of Probability Generating Functions"**, *J. Inst. Maths. Applics.*, Vol 22, pp. 275-282, 1978.

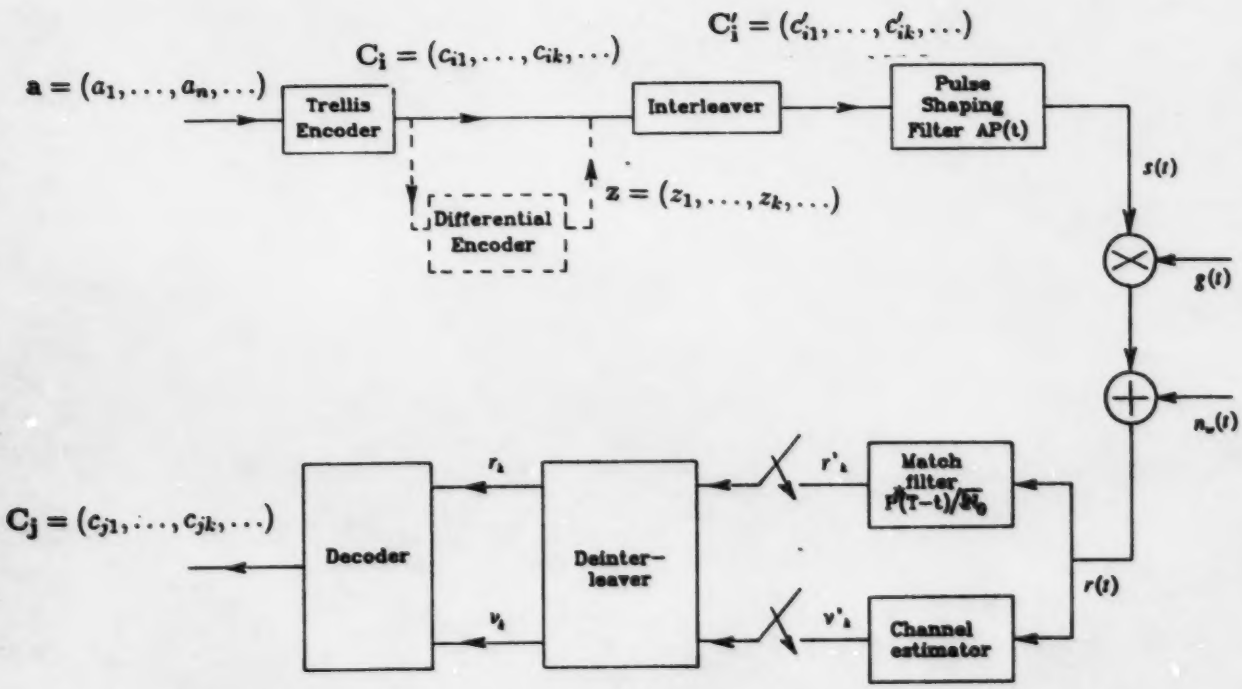

Figure 1: Block diagram of the trellis coded system

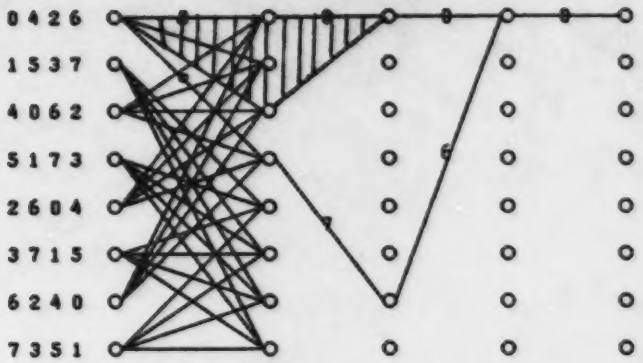

Figure 2: Ungerboeck's 8 state 8PSK code

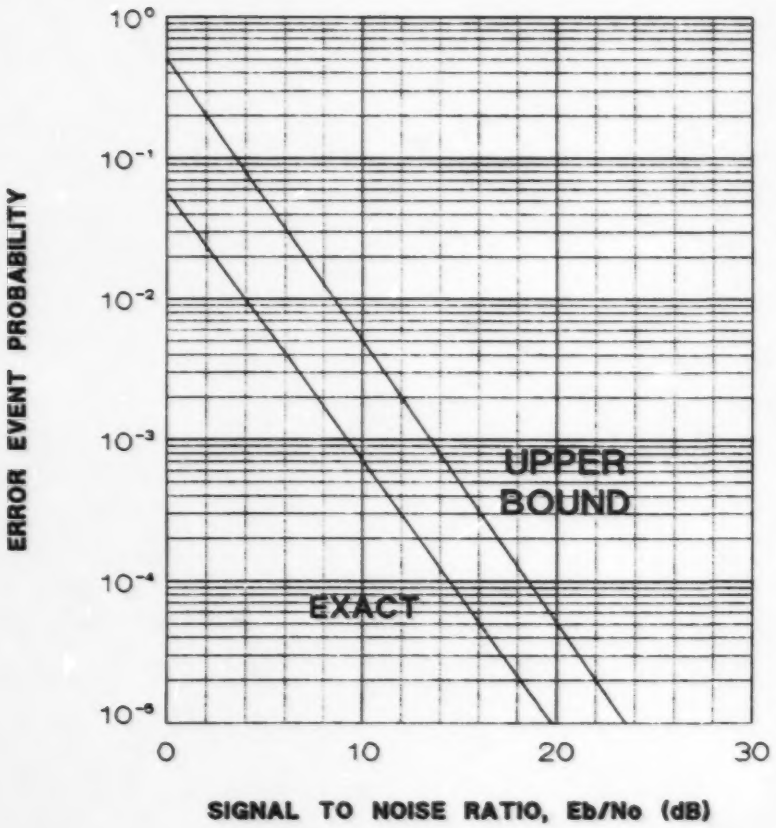

Figure 3: Exact and upperbound on the probability of the shortest error event of Ungerboeck's 8 state 8PSK code in a Rayleigh fading channel with ideal interleaving and perfect CSI

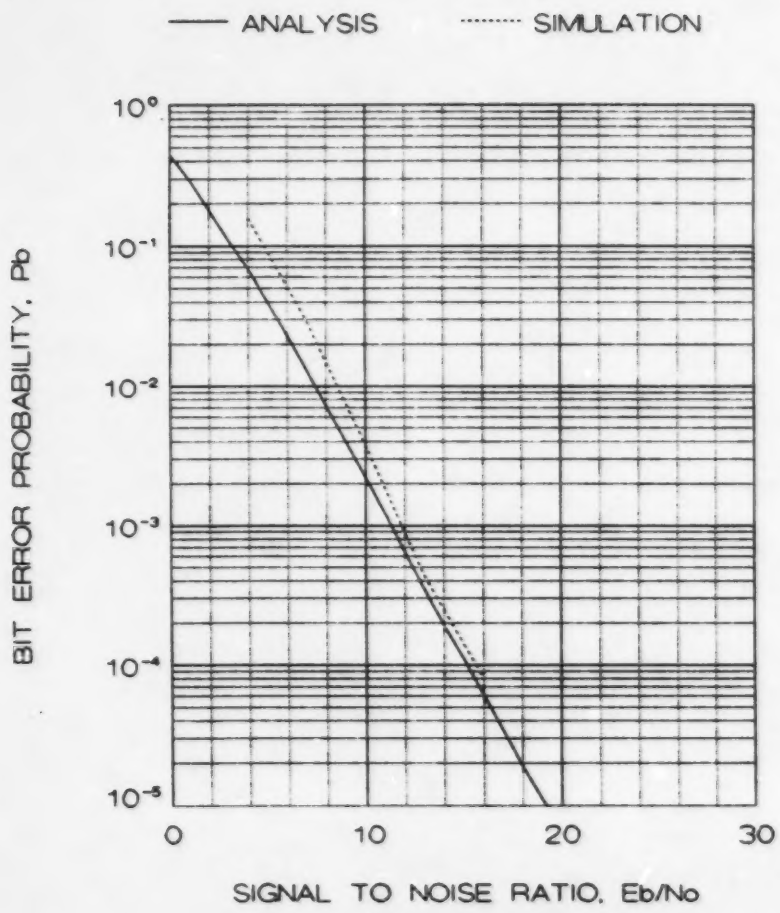

Figure 4: Bit error probability of Ungerboek's 8 state 8PSK code in a Rayleigh channel with ideal interleaving and perfect CSI.

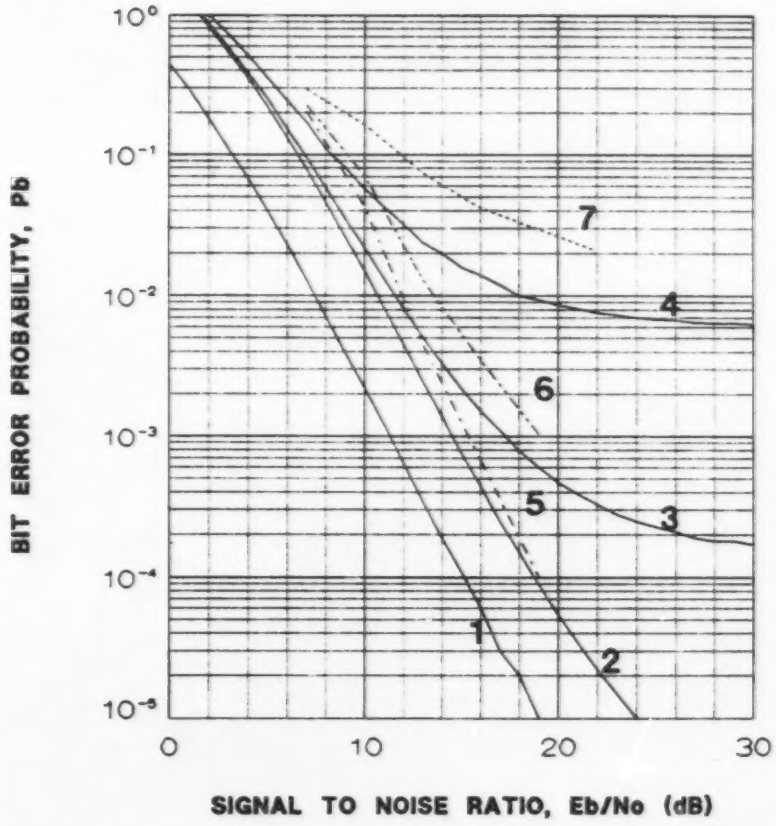

Figure 5: Bit error probability of Ungerboeck's 8 state 8PSK code in a Rayleigh channel with differential detection and ideal interleaving. Curves 2,3, and 4 are the analytical results while curves 5,6,and 7 are the corresponding simulation results. The fade rate in curves 2 and 5, 3 and 6, and 4 and 7, are .01, .03, and .06 respectively. Curve 1 is the analytical result from Fig. 4.

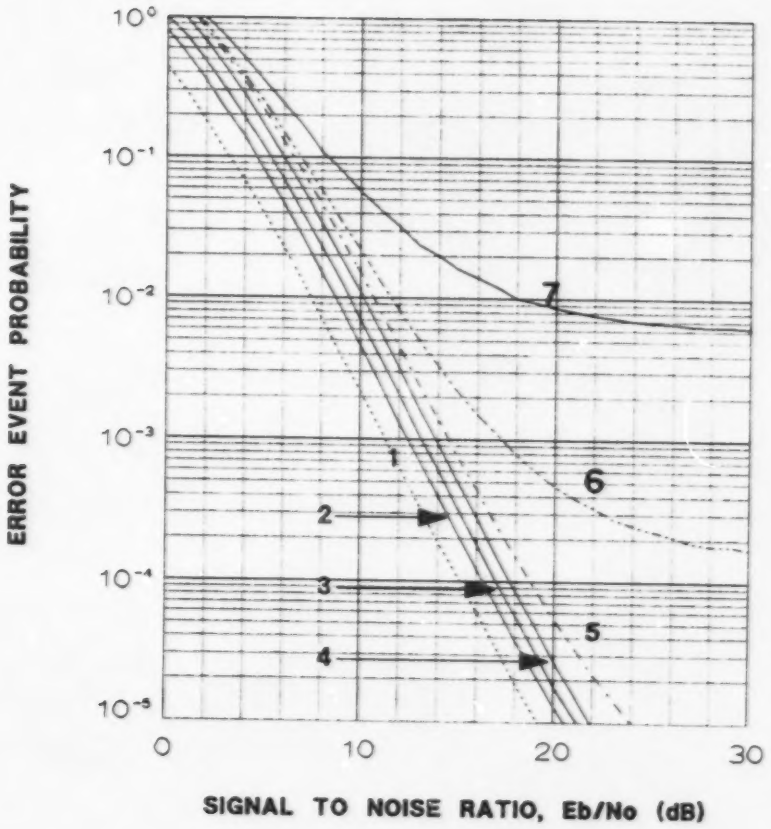

Figure 6: Bit error probability of Ungerboeck's 8 state 8 PSK code in a Rayleigh channel with ideal interleaving and pilot tone aided detection. Curve 1 is the idealized result from Fig. 5, curves 2, 3, and 4 are the results for pilot-aided detection and fade rates of .01, .03, and .06 respectively, curves 5, 6, and 7 are the DPSK results from Fig. 5.

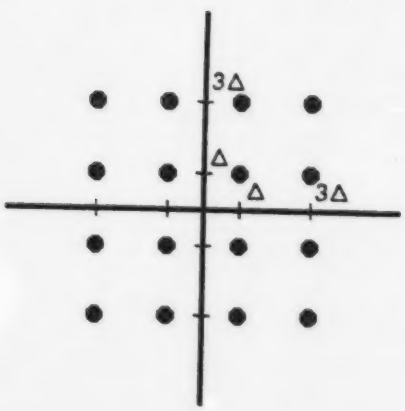

Figure 7: Signal constellation of 16 QAM

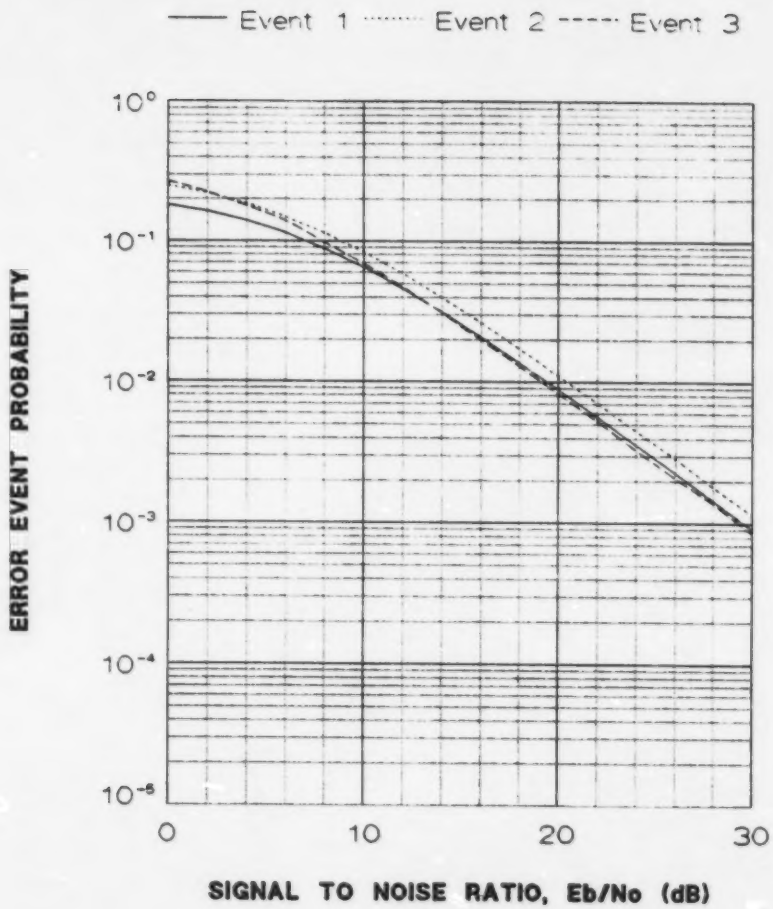

Figure 8: Error event probability of 16 QAM in a Rayleigh channel with ideal interleaving and pilot-tone aided coherent detection. The bandwidth of the pilot filter is 0.12. Error event 1 is the event of confusing the transmitted symbol $(\Delta, 3\Delta)$ with the symbol $(3\Delta, 3\Delta)$, Event 2 is the event of confusing $(3\Delta, 3\Delta)$ with $(\Delta, 3\Delta)$, and Event 3 is the event of confusing (Δ, Δ) with $(\Delta, -\Delta)$.

Centre for Systems Science

• Simon Fraser University established the **Centre for Systems Science** as a provincial centre for advanced study. The objective of CSS is twofold: coordinate and combine advanced systems research at SFU, and encourage application of research results within industry in order to further the technological development of British Columbia.

• CSS has assembled three research groups to streamline its efforts. The orientation of these groups is as follows: intelligent systems, microelectronics, and computer and communications systems. Because of shared research interests, this latter group merges with LCCR (**Laboratory for Computer and Communications Research**) for the publication of technical reports.

• CSS's three groups involve over 80 skilled faculty researchers principally from SFU's Schools of Engineering Science and Computing Science, but also draw substantially on Mathematics, Physics, Geography, Philosophy, Linguistics, Psychology, Communication, and Kinesiology.